

Příslovečné spřežky, slovnědruhové přesahy tvary na o

PLIN032

O čem budeme hovořit

- Adverbium jako slovní druh
- Jak jsou lemmatizovány a slovnědruhově interpretovány tvary adverbií, které končí na –o v korpusech SYN
- Jak vyhledat tvary potenciálních adverbiálních spřežky i nespřážené ustrnulé tvary (nikoli jen po předložce)
- Návrh lemmatizace

Adverbium jako slovní druh

- Neohebný, autosémantický, primární syntaktická funkce adverbiální určení
- Neohebnost a původ adverbií – slovnědruhová charakteristika
- Slovnědruhová přesažnost zájmenných příslovcí
- Adverbializace jmen – adverbiální spřežky
- Kodifikace variant

Slovnědruhová interpretace adverbii

- Dg.*
- Db.*
- X.*
- Jiné

Jak jsou lemmatizovány a slovnědruhově interpretovány tvary adverbii, které končí na **-o** v korpusech SYN

- [lc=".*o" & tag="D.*"]

Frekvenční distribuce x

[Běžná](#) [Podle typů textů](#) [Závislost dvou atributů](#)

Frekvenční limit:

Úroveň	Atribut	Nerozlišovat velikost	Pozice	(Node) začít od	
1.	lc [lowercase word]	<input type="checkbox"/>	Node	slova KWIC nejvíce vlevo	▲ x
2.	lemma_lc [lowercase lemma]	<input type="checkbox"/>	Node	slova KWIC nejvíce vlevo	▲ x
3.	tag	<input type="checkbox"/>	Node	slova KWIC nejvíce vlevo	▲ x

[+](#)

[Vytvořit frekvenční seznam](#)

Minimální frekvence:

[Použít](#)

Celkem: 373 položek (8 stránek)

	Filter	lc [lowercase word]	lemma_lc [lowercase lemma]	tag	Freq	
1	p / n	proto	proto	Db-----	64168	
2	p / n	často	často	Dg----1A---	43788	
3	p / n	co	co	Db-----	34447	
4	p / n	přímo	přímo	Dg----1A---	30191	
5	p / n	dlouho	dlouho	Dg----1A---	29615	
6	p / n	přesto	přesto	Db-----	24591	
7	p / n	skoro	skoro	Db-----	22441	
8	p / n	daleko	daleko	Dg----1A---	16710	
9	p / n	málo	málo	Dg----1A---	15980	
10	p / n	ráno	ráno	Db-----	15388	
11	p / n	naprosto	naprosto	Db-----	14021	
12	p / n	snadno	snadno	Dg----1A---	11728	
13	p / n	dávno	dávno	Dg----1A---	11161	
14	p / n	takto	takto	Db-----	11146	
15	p / n	těžko	těžko	Dg----1A---	9222	
16	p / n	nedávno	nedávno	Db-----	8585	
17	p / n	najevo	najevo	Db-----	5140	
18	p / n	vlevo	vlevo	Db-----	4157	
19	p / n	vpravo	vpravo	Db-----	3957	
20	p / n	líto	líto	Db-----	3949	
21	p / n	blízko	blízko	Dg----1A---	3671	
22	p / n	vysoko	vysoko	Dg----1A---	3654	
23	p / n	hluboko	hluboko	Dg----1A---	3653	
24	p / n	nato	nato	Db-----	3285	

Otázky

- Proč jsou některá označována jako stupňovatelná a jiná nikoli?
- Je pozitivem od *snadněji* vždy tvar *snadno*?
- Proč je lemmatem tvaru *brzo* tvar *brzy*? Existuje více takových případů? Jak formulovat dotaz, abychom podobné případy našli?


P-filtr [lc=". *o" & lemma!=". *o"]

Frekvenční seznam

1 / 1

Minimální frekvence: [Použít](#)

Celkem: 7 položek (1 stránka)

	Filter	lc [lowercase word]	lemma_lc [lowercase lemma]	tag	Freq	
1	p / n	brzo	brzy	Dg-----1A--1-	2203	
2	p / n	blizoučko	blizoučce	Db-----	48	
3	p / n	teďko	teďka	Db-----6-	21	
4	p / n	stopro	stoprocentně	Dg-----1A---	19	
5	p / n	ko	ko	Db-----8-	3	
6	p / n	al-ko	al-ko	Db-----8-	1	
7	p / n	dto	dto	Db-----8-	1	

hledání dvojic nástrojem Morfio

Morfio

<+ společný odlišný +> Morf. specifikace:

vzor 1:

vzor 2:

Další vzor

Korpus: Frekvence vyšší než: Hledají se: Vyhodnocují se:

Velikost písmen: ignorovat

Odkaz na toto zadání: <http://morfio.korpus.cz/SFr2qjdl>

výpis

vzor 1 ▲▼ (fq ▲▼)	vzor 2 ▲▼ (fq ▲▼)
1 brzy (16127)	brzo (2468)
2 ted'ky (3)	ted'ko (59)

Jak vyhledat tvary potenciálních adverbálních spřežky i nespřažené ustrnulé tvary (nikoli jen po předložce)

- Podívejte se do korpusu na tvary předložkových spojení a spřežek.
- Vyhledejte v korpusu dvojice adverbií od stejného základu lišící se slovnědruhovou charakteristikou (např. *smutně/smutno*).

dvojice [eě]/o

Pozorujte vyhledané dvojice a pokuste se je klasifikovat.



<+ společný odlišný >+ Morf. specifikace:

vzor 1:

vzor 2:

Další vzor

Korpus: Frekvence vyšší než: Hledají se: Vyhodnocují se:

Velikost písmen: ignorovat

Odkaz na toto zadání: <http://morfio.korpus.cz/AlwRKQRk>

Zopakujte vyhledávání tak, aby typ na o nebyl rozpoznán automatickou analýzou.

<+ společný odlišný +> Morf. specifikace:

vzor 1:

vzor 2:

Další vzor

Korpus: Frekvence vyšší než: Hledají se: Vyhodnocují se:

Velikost písmen: ignorovat

Odkaz na toto zadání: <http://morfio.korpus.cz/WKc9fFMJ>

Úkoly

- **Výsledky seřadíte podle frekvence vzoru2.**
- **Definujte nedostatky slovníku použitého analyzátoru.**
- **Podívejte se na typ adverbíí na *y*. Jak jsou lemmatizovány a značkovány tvary typu *po .*[sc]ku*?**

Jak lemmatizovat a interpretovat tvar *such*, *sucha*, *sucho*, *suchy* ale i *suší*?

- a. *Bylo velké sucho.* (substantivum → POS=N)
- b. *Bylo velmi **sucho**.* (adverbium → POS=D);
- c. *Bylo **sucho**.* (adverbium → POS=D)
- d. *Dítě bylo **sucho**.* (adjektivum – jmenný tvar → POS=A);
- e. *Chlapi seděli v hospodě **na sucho**.* (adverbium → POS=D)
- f. *Chlapi si v hospodě stěžovali **na sucho**.* (substantivum → POS=N)
- g. ***Do sucha**, které ničilo úrodu v roce 1947, jsem hroznější počasí nezažil.* (substantivum → POS=N)
- h. *Utři se **do sucha**.* (adverbium → POS=D)
- i. *Jsem už zcela **such**, manželka je **sucha** i dítěta jsou již **sucha**, ale šaty nejsou dosud zcela **suchy**.* (adjektivum – jmenný tvar → POS=A)
- j. *Tolik **such** nebylo ani v minulém století.* (substantivum → POS=N)
- k. *Tak to by bylo v **suchu**.* (adverbium → POS=D)
- l. *V takovém dlouhotrvajícím **suchu** dojde snadno k požáru.* (substantivum → POS=N)
- m. *Nevěděli jsme si rady s velkými **suchy**.* (substantivum → POS=N)

Závěr

- **Jmenné tvary adjektiv**
- **Adverbializace jmenných tvarů**
- **Pravidla pro zaznamenávání adverbializovaných jmenných tvarů ve slovnících automatických analyzátorů**
- **Pravidla pro desambiguaci homonymní tvarů**

Úkol na 18. 11. 2020

- **V návrhu na nové značkování české morfologie (NovaMorf) se předpokládá, že tvary pasivních příčestí budou lemmatizovány několikanásobným lemmatem, a sice dlouhým tvarem adjektiva tvořeného z krátkého tvaru (příčestí) a infinitivem.**
- **Jednou z námitek proti tomuto řešení byla údajná nedoloženost dlouhých tvarů.**
- **V synv7 jsou doloženy tvary pasiv bez odpovídajících složených tvarů:**
*důvěřováno, křivděno,
reptáno, hřešeno, poshověno, hýbáno, naléháno, přáno, zacházeno,
pohlíženo, jednáno, vyhrožováno, hovořeno, přihlíženo, hleděno,
zabraňováno, vycházeno, pracováno, referováno, reagováno, dohlíženo,
pátráno... a další.*
- **Projděte korpusy (webové – řada Araneum, czTenTen) a shromážděte doklady odporující uvedené námitce. Pokuste se najít důvody pro výskyt doložených dlouhých tvarů.**